J. Adelaide Bot. Gard. 7(2): 159-166 (1985)

AN OBJECTIVE METHOD FOR ASSESSING THE PERFORMANCE OF AMENITY PLANTINGS

Brendan G. Lay and Anthony P. Meissner

S.A. Department of Agriculture, 25 Grenfell Street, Adelaide, South Australia, 5000

Abstract

This paper describes the conceptual development, use and statistical validation of an objective method by which the performance of broadscale shrub and tree plantings for amenity purposes can be determined. The method provides an evaluation of the performance of specimens in relation to their potential on a species basis. It gives consistent results on the relative performance of species planted at a particular site and under particular horticultural treatments. The paper also highlights the need to statistically test the repeatability of subjective assessments used in other amenity plant surveys, particularly those used for monetary valuations.

Introduction

Amenity can be defined as "the subjective human assessment of pleasantness". As such, it is easy to perceive why those attempting to put amenity values on horticultural plantings have experienced so many problems in producing a universally acceptable and repeatable method. Many attempts have also been made to develop methods for determining tree value in monetary terms; these have been reviewed by Raad (1972) and Wycherley (1979). The primary aim of all these methods was to establish an objective basis for the assessment of amenity value in monetary terms of individual trees or shrubs to people living in the locality. The approach taken in these methods has been to determine the size and condition of the tree, then assess the significance of its location and finally to combine, usually by multiplication, these factors with a money index to determine a final monetary value. The methods, in effect, attempted to provide objectively based criteria to define a subjective concept. Since all the methods were trying to achieve the same goal, viz. the determination of a realistic monetary value of an individual plant which can stand up in legal proceedings, then one method should be agreed to and adopted by all responsible bodies in Australia. However, this is not the case. At present there are at least three approaches being used in different parts of Australia all based on techniques developed overseas. These methods cover trees only and are described by Boomsma (1973), Australian Institute of Horticulture (1977) and Kartzoff (1977).

There are two serious weaknesses in all the currently used methods of amenity tree valuation: (a) the final monetary value is derived by multiplying a number of factors together and hence an error in one factor is compounded in the calculation and (b) there has been no attempt to determine repeatability of any method by different observers. This need has been recognised. For example the Australian Institute of Horticulture Committee report (1977) states that the elimination of bias ". . . is regarded as most important, particularly to ensure some uniformity of judgments from different valuers and to minimize the spread of subjectively assessed results". Despite this, to our knowledge, no evaluation of assessment methods, statistical or otherwise, has been published.

This paper describes an objective method of assessing performance of amenity plantings. The individual trees or shrubs are assessed and scored with reference to optimum observed performance for that species in cultivation under field conditions. No reference is made to significance of location or monetary value. As such it avoids most of the subjectivity of the above methods. In this respect it is similar to methods used in some overseas surveys. For example Good *et al.* (1978) in a broadscale survey of amenity trees, classified plants on a species basis but made no attempt to determine their condition.

Developing the method

The need to develop a method for objectively assessing amenity plantings arose in 1979 when we were asked to evaluate the performance of individual plants at a variety of locations in South Australia. It soon became obvious that evaluation of the elusive characteristic "amenity value" could not be done by physical measurement alone. For example, a healthy slow growing specimen should score more highly than one which has grown faster but was subject to wind damage, was diseased, or looks unsightly. In addition, any useful method had to be practical, because individual plant ratings needed to be performed rapidly. At one site, for example, about 12,000 trees were assessed. Physical measurements would have been too time consuming.

The method as conceived uses two criteria, health and vigour, which together indicates the performance of trees or shrubs anywhere at any time. Health is assessed by rating the plant on a scale of 0-5 relative to a completely healthy and undamaged specimen, while vigour, or rate of growth of the plant is also scored on a similar scale with reference to maximum expected growth for that species under cultivation. This requires that observers should have some knowledge of expected growth rates of individual species at a given age. Observers who are unfamiliar with the species assessed can gain a reasonable idea of anticipated growth rates by referring to the species or taxon in published works, in botanical gardens, or by comparison with other plantings of known age where that taxon is represented. The optimum growth rates refer to the best growth commonly seen in cultivation, or locally observed, rather than to the optimum in its natural environment.

Trunk form as used in forestry assessments was not considered relevant as it is unrelated to amenity value. However, height and girth measurements can be correlated with vigour ratings for estimates of timber volume.

Table 1. Criteria used in rating plants for amenity purposes.

(a) HEALTH SCALE (0-5)

0.			Plant dead
1.		(a)	No foliage, stems still green
	or	(b)	100% dieback or supression of terminal foliage, any new growth or resprouts unhealthy, chlorotic or absent.
2.		(a)	100% dieback or supression of terminal foliage, new growth or resprouts healthy.
	or	(b)	Apparently chronic or systemic infection or dessication with 75-100% of foliage dead, lost or damaged.
	or	(c)	Two or more of the factors under "3" below.
3.		(a)	Whole plant showing chlorosis, including new growth.
	or	(b)	Most leaves lost on lower growth, healthy tip growth remaining.
	or	(c)	50-70% of foliage affected by disease and/or dessication.
	or	(d)	Death or dieback of a major stem or portion of canopy; remainder healthy.
4.		(a)	Healthy plant but with significant (25-50%) leaves lost or damaged.
	or	(b)	Healthy, with minor stem or canopy damage (affecting less than 25% of plant).
	or	(c)	Chlorosis of non-terminal foliage.
	or	(d)	Slight ill-thrift generally apparent.
5.			Healthy, but includes plants with up to 25% of leaves damaged in some way.

Table 1 *(continued)* **(b) VIGOUR SCALE (0-5)**

0.			Plant dead.
1.		(a)	No recent increase in canopy; size less than 25% of optimum.
	or	(b)	New growth, but plant less than 10% of optimum.
2.		(a)	Growth less than 25% of optimum, new leaves but only slight recent increase in canopy size.
	or	(b)	Growth less than 25% of optimum, major stem resprouting.
3.			Growth 25-50% of optimum.
4.			Growth 50-75% of optimum.
5.			Growth 75-100% of optimum.

The method was field tested and modified to cater for most field situations. The final list of criteria is shown in Table 1. As can be seen from this table, the intention was to cover all commonly observed variations in plant performance with no direct measurements necessary. Where percentages are mentioned, they refer to linear proportions. For example, a tree growing at 50% of the optimum is, on average, half the height and half the canopy diameter of the optimum standard. A plant of maximum expected canopy diameter but half the optimum height is assessed at 75% of the optimum under the vigour score criteria. That is, an average is taken of the linear measurements. This approach simplifies field ratings.

No stem dimensions are used in the method, because of the difficulty of rapid measurement and the need to embrace multi-stemmed trees, shrubs and ground-cover plants in the assessment. The need for observers to have prior knowledge of species characteristics seemed desirable, if not essential, so this requirement was tested using a student from Roseworthy Agricultural College who was inexperienced in plant assessment.

The method was first used at the Monarto Irrigation Experiment Station near Murray Bridge (35° 10′S 139° 17′E) and later at the Arid Zone towns of Woomera, (31° 11′S 136° 54′E), Leigh Creek (30° 31′S 138° 25′E) and Radium Hill (32° 30′S 140° 32′E), South Australia. At Monarto it proved effective in distinguishing the effects of irrigation treatments and soil type on trees, shrubs and groundcovers (Lay, 1980). In this example, scores were averaged for all 46 species to give relative plant performance.

Statistical analysis

The analysis used data from field tests at Monarto and Woomera. The data consisted of four sets of observations. These were (1) Two trained observers; one experienced in plant assessment, the other relatively inexperienced, independently assessed 460 trees and shrubs at Monarto in March 1980. (2) An experienced observer rated 46 plants at Monarto in July 1982 and repeated it a week later. (3) A set of observations (992) was taken in May 1983 by an experienced observer followed a month later by an inexperienced observer. (4) Three trained observers rated a group of trees, either singly or as a clump (27 observations) at Woomera in November 1982. In each case the observers independently scored individual plants on the health and vigour scale as outlined in Table 1.

Contingency tables of agreement between observers were drawn up from the health and vigour scores (see Appendix). The method of analysis tested whether observers were ranking the plants similarly. If the observers agreed on the ranking then the tables would be symmetrical, but if they disagreed then non-symmetrical tables would result. The tables were said to be quasi-symmetrical if all the values for a particular row were a multiple of the corresponding column. (Bishop, Fienberg and Holland, 1975). That is, observers applied a constant bias to each category independent of each other. A significant test value (chi-squared) indicated this condition. If this was not so, then the tables were tested for overall symmetry which was indicated by a non-significant test value. If the tables were symmetrical overall then this implied that the row and column totals were also equal (marginal homogeneity).

Table 2. Results of statistical analyses for data sets 1-4. Figures are chi-squared values.

Data set		Health		Vigour	
1	Quasi-symmetry	4.828	n.s.	2.241	n.s.
	Symmetry	29.41	*	8.823	n.s.
	Marg. Homegeneity	24.58	***	6.582	n.s.
2	Quasi-symmetry	0.0003	n.s.	0.0003	n.s.
	Symmetry	4.271	n.s.	14.29	n.s.
	Marg. Homogeneity	4.271	n.s.	14.29	n.s.
3	Quasi-symmetry	20.47	*	27.41	***
4	Quasi-symmetry	10.03	*	0.0002	n.s.
obs.	Symmetry	—		4.159	n.s.
1 vs 2	Marg. Homogeneity	—		4.159	n.s.
	Quasi-symmetry	0.0001	n.s.	0.0001	n.s.
obs.	Symmetry	7.070	n.s.	12.68	*
1 vs 3	Marg. Homogeneity	7.070	n.s.	12.68	**
	Quasi-symmetry	0.0001	n.s.	0.0001	n.s.
obs.	Symmetry	12.82	*	12.14	n.s.
2 vs 3	Marg. Homogeneity	12.82	**	12.14	**

n.s. Not significant   * Significant at 5% level
** Significant at 1% level   *** Significant at 0.1% level

Results and discussion

(a) Data analyses

The data sets are presented as contingency tables of health and vigour scores in the Appendix. Results of the statistical analyses are shown in Table 2. The tables were symmetrical for both health and vigour only for data set 2 (self-consistency test) i.e. the observer was consistent in his rating ability even though a week had elapsed between observations. Observer bias was indicated by the significant chi-squared value for data set 3 (inexperienced observer) where the model of quasi-symmetry was rejected for both health and vigour. This may have been due to some mis-identification of plants by the Roseworthy student as indicated by the ratings of 4 and 5 against ratings of 0 and 1 by the other observer and vice versa in the table. The results for data set 4 were variable possibly because of the low total numbers of observations.

The statistical evaluation does highlight several matters in regard to the method. Firstly, over a short time span an observer can consistently give an equal rating to the same set of plants. Hence, there should be little difference in day to day ratings when assessing large amenity plantings. Secondly, any method of visual assessment, no matter how inherently objective it is, requires training of the observers. This is shown particularly by the results of the inexperienced Roseworthy student (data set 3) and to a lesser extent by the tables which displayed asymmetry. Discrepancies in ratings will always occur between observers because of their different perceptive abilities and the weight each of them places on the method's criteria. It is therefore essential for observers to spend time together on calibration and where possible, to statistically analyse the observations taken to check for bias. The calibration should cover the full range of the rating scale.

The analyses overall indicate that criteria derived for distinguishing between the different scores can lead to objective assessment of plants whether they are trees, shrubs or groundcover. This contention is further borne out if the degree of agreement between observers is considered. Complete agreement amongst observers, (the sum of the main diagonal of the table), ranged from 44% (data set 4, observer 1 vs observer 2, health) to 79% (data set 1, health). The diagonals on either side of the main one represent a difference in rating of one point between observers. The sum of these "one-off" values plus the values for complete agreement, ranged from 89% (data set 3, vigour) to 100%. Hence, even though there was significant observer bias in some instances, they agreed within one point of each other in more than 90% of cases.

(b) Practical applications

The method was developed to determine the relative performance of a range of plants at a given locality or under different cultural conditions. The individual scores for the two categories can be used in two ways. Firstly, where the aim of the exercise is to provide a ranked performance summary, then all that is required is the overall rating value, produced by summing the individual health and vigour assessments. Averaging the individual ratings for each species can then provide an average performance score on a 10 point scale. This score can then provide a basis for simple and commonly accepted performance categories according to the following table:

Average rating (Health + vigour)	Performance Category
8 - 10	Satisfactory/recommended
6 - 7	Indeterminate
0 - 5	Poor/not recommended

This method has proved adaptable and for most evaluations of plantings the above divisions have been satisfactory. The method was used recently by Bulman (1983) and Lay (1980, 1983). Moreover it is obvious that the more harsh the site is, the lower the admissible performance scores for recommended species.

Secondly, if a more detailed analysis of environmental or treatment effects on growth is required, then it is not desirable to combine the individual health and vigour scores in an overall rating value. For example, variation in performance of some plants may be due to different growth rates or to different degress of susceptibility to leaf-chewing insects. There is a need under some circumstances, e.g. provenance trials, to evaluate factors affecting performance of a taxon at a specific or even subspecific level. It is more important for amenity purposes that a plant looks healthy though it may grow slowly, than a faster growing specimen which looks unhealthy much of the time due to damage or disease.

The main limitation of the method as an objective and scientific approach to plant assessment is that observers need to know potential growth rates and final sizes of species. In practice, however, this limitation is not as serious as it may first appear for the following reasons:

a) Final size of most species under cultivation can be obtained from published works or from observations of old established plantings. These sizes can be recorded and referred to in the field.

b) The vigour scales are not sensitive to large differences in expected optimal growth as perceived by different observers. As can be seen from Table 1 the vigour criteria uses the average of the linear measurements (height and canopy dimensions). If the optimal height and canopy dimensions as perceived by one observer are half of those as discerned by another, a difference in one point arises. Hence a large variation between observers in expected growth will only have a small effect on the overall rating given.

c) Many amenity surveys require relative performance of different taxa at a site or between different sites and so absolute values are not important.

Errors are more likely to occur in young plantings where less information is usually available to both experienced and inexperienced observers on a growth-for-age basis. However, any performance figures must be considered tentative only where plantings are young (see comments in Lay, 1983). It is the relative performance of taxa in these young plantings which is more important.

Conclusions

This objective approach to the evaluation of amenity plantings provides a statistically repeatable basis for determining the subjective concept of amenity value. We believe that it can provide a meaningful and objective basis to the assessment of horticultural performance of various plant taxa. In addition, when incorporated with other criteria, it can also be used to more objectively define monetary value of individual plant specimens.

Acknowledgements

We are grateful for the considerable time spent by our observers and colleagues in developing the method and survey work—Mr B. Dubois and Mr E. Thexton at Monarto and Mr A. Beal and Mr J. Zwar at Woomera. We also thank Mr P. McCloud for statistical advice on the analysis of contingency tables.

References

Australian Institute of Horticulture (1977). A method for assessing the monetary value of trees used in amenity horticulture. "Technical memorandum No. 1", July 1977, 14pp.

Bishop, Y.M.M., Fienberg, S.E. and Holland, P.W. (1975). Analysis of Square Tables: Symmetry and Marginal Homogeneity pp 281-310 In "Discrete Multivariate Analysis: Theory and Practice". (M.I.T. Press: Cambridge, U.S.A.).

Boomsma, C.D. (1973). The Monetary Value of Amenity Trees in Public Places. *Australian Parks* 36-39.

Bulman, P.U. (1983). Broadscale Planting Techniques and Species evaluation at Monarto. "S.G.A.P. Seminar papers, Lincoln College". (S.G.A.P.: Adelaide).

Good, J.E., Craigie, I., Last, F.T. and Munro, R.C. (1978). Conservation of Amenity Trees in the Lothian Region of Scotland. *Biological Conservation* 13 (4) 247-272.

Kartzoff, O.A.M. (1977). Amenity tree valuation for Australia. *Institute of Foresters of Australia, Newsletter.* 18 (3): 24-27.

Lay, B.G. (1980). Assessment of Amenity Tree planting on the Monarto Irrigation Experiment Station. "South Australian Department of Agriculture, Soil Conservation Branch Report S7/80", 15pp.

Lay, B.G. (1983). Effects of Irrigation Regimes on Performance of Native Plants Under Cultivation pp 31-43 In "Biology of Native Plants Under Cultivation". (South Australian Department of Environment and Planning: Adelaide).

Raad, A. (1972). The Value of Trees. *Australian Parks* 8 (4) 507.

Wycherley, P.R. (1979). Amenity Tree Valuation—A Working Paper. *Supplement to Australian Parks and Recreation*, 15, 9 pp.

Appendix

Data set 1

Health, Observer 1

	Scores	0	1	2	3	4	5	Marg. total
Observer 2	0	30	1	0	0	0	0	31
	1	0	10	2	0	0	0	12
	2	0	4	8	3	1	0	16
	3	0	3	3	37	9	0	52
	4	0	0	1	25	71	49	146
	5	0	0	0	2	20	181	203
Marg. Total		30	18	14	67	101	230	460

Vigour, Observer 1

	Scores	0	1	2	3	4	5	Marg. total
Observer 2	0	30	1	0	0	0	0	31
	1	0	31	4	1	0	0	36
	2	0	6	30	20	1	0	57
	3	0	0	19	53	20	1	93
	4	0	0	2	20	92	26	140
	5	0	0	0	1	12	90	103
Marg. Total		30	38	55	95	125	117	460

Data set 2

Health, Observer 1

	Scores	0	1	2	3	4	5	Marg. total
Observer 2	0	3	0	0	0	0	0	3
	1	0	1	2	0	0	0	3
	2	0	0	0	1	0	0	1
	3	0	0	0	4	1	0	5
	4	0	0	0	1	3	4	8
	5	0	0	0	0	5	21	26
Marg. Total		3	1	2	6	8	25	46

Vigour, Observer 1

	Scores	0	1	2	3	4	5	Marg. total
Observer 2	0	3	0	0	0	0	0	3
	1	0	2	0	0	0	0	2
	2	0	3	0	1	0	0	4
	3	0	0	1	2	2	0	5
	4	0	0	0	0	4	9	13
	5	0	0	0	0	1	18	19
		3	5	1	3	7	27	46

Data set 3

Health, Observer 1

	Scores	0	1	2	3	4	5	Marg. Total
Observer 2	0	182	0	0	1	1	2	186
	1	0	5	13	2	0	4	24
	2	0	4	35	15	8	5	67
	3	0	7	17	38	41	11	114
	4	0	3	12	49	113	94	271
	5	2	0	4	22	109	193	330
Marg. Total		184	19	81	127	272	309	992

Vigour, Observer 1

	Scores	0	1	2	3	4	5	Marg. Total
Observer 2	0	182	0	0	0	2	2	186
	1	1	14	30	11	6	3	65
	2	1	0	33	23	13	11	81
	3	0	0	33	80	50	20	183
	4	0	0	9	68	120	42	239
	5	0	1	4	26	101	106	238
Marg. Total		184	15	109	208	292	184	992

Health
Observer 1

	Scores	(0,1,2)	3	4	5	Marg. Total
Observer 2	(0,1,2)	1	0	1	0	2
	3	4	4	0	0	8
	4	0	6	2	0	8
	5	0	0	4	5	9
Marg.	Total	5	10	7	5	27

Health
Observer 1

	Scores	(0,1,2)	3	4	5	Marg. Total
Observer 3	(0,1,2)	5	2	1	0	8
	3	0	3	1	0	4
	4	0	5	3	2	10
	5	0	0	2	3	5
Marg.	Total	5	10	7	5	27

Health
Observer 2

	Scores	(0,1,2)	3	4	5	Marg. Total
Observer 3	(0,1,2)	2	5	1	0	8
	3	0	2	2	0	4
	4	0	1	6	3	10
	5	0	0	0	5	5
Marg.	Total	2	8	9	8	27

<u>Data set 4</u>

Vigour
Observer 1

	Scores	(0,1,2)	3	4	5	Marg. Total
Observer 2	(0,1,2)	1	2	0	0	3
	3	0	7	1	0	8
	4	0	3	7	1	11
	5	0	0	2	3	5
Marg.	Total	1	12	10	4	27

Vigour
Observer 1

	Scores	(0,1,2)	3	4	5	Marg. Total
Observer 3	(0,1,2)	3	5	1	0	9
	3	0	1	3	0	4
	4	0	2	7	3	12
	5	0	0	0	2	2
Marg.	Total	3	8	11	5	27

Vigour
Observer 2

	Scores	(0,1,2)	3	4	5	Marg. Total
Observer 3	(0,1,2)	1	8	0	0	9
	3	0	3	1	0	4
	4	0	1	8	3	12
	5	0	0	1	1	2
Marg.	Total	1	12	10	4	27